Fuzzy and Rough Set Approaches for Uncertainty in Spatial Data

Theresa Beaubouef and Frederick E. Petry

Abstract. The management of uncertainty in databases is necessary for real world applications, especially for systems involving spatial data such as geographic information systems. Rough and fuzzy sets are important techniques that can be used in various ways for modeling uncertainty in data and in spatial relationships between data entities. This chapter discusses various approaches involving rough and fuzzy sets for spatial database applications such as GIS.

Keywords: Spatial Interpolation, Triangulated Irregular Networks, Spatial Data Mining, Minimum Bounding Rectangles, Rough Sets, Topological Spatial Relations.

1 Introduction

A spatial database is a collection of data concerning objects located in some reference space, which attempts to model some enterprise in the real world. The real world abounds in uncertainty, and any attempt to model aspects of the world should include some mechanism for incorporating uncertainty. There may be uncertainty in the understanding of the enterprise or in the quality or meaning of the data. There may be uncertainty in the model, which leads to uncertainty in entities or the attributes describing them. And at a higher level, there may be uncertainty about the level of uncertainty prevalent in the various aspects of the database. There has been a strong demand to provide approaches that deal with inaccuracy and uncertainty in geographical information systems (GIS) and their underlying

---

Theresa Beaubouef
Department of Computer Science and Industrial Technology
Southeastern Louisiana University
Hammond, LA 70402

Frederick E. Petry
Mapping Charting and Geodesy Branch
Naval Research Laboratory
Stennis Space Center, MS 39529

R. Jeansoulin et al. (Eds.): Methods for Handling Imperfect Spatial Info., STUDFUZZ 256, pp. 103–129.
springerlink.com

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**
4. TITLE AND SUBTITLE **Fuzzy and Rough Set Approaches for Uncertainty in Spatial Data**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,Mapping Charting and Geodesy Branch,Stennis Space Center,MS,39529**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The management of uncertainty in databases is necessary for real world applications, especially for systems involving spatial data such as geographic information systems. Rough and fuzzy sets are important techniques that can be used in various ways for modeling uncertainty in data and in spatial relationships between data entities. This chapter discusses various approaches involving rough and fuzzy sets for spatial database applications such as GIS.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**28**	

spatial databases. The issue of spatial database accuracy has been viewed as critical to the successful implementation and long-term viability of GIS technology [63]. There are a variety of aspects of potential errors in GIS encompassed by the general term "accuracy." However, here we are only interested in those aspects that lend themselves to modeling by fuzzy and rough set techniques.

Many operations are applied to spatial data under the assumption that features, attributes and their relationships have been specified a priori in a precise and exact manner. However, inexactness often exists in the positions of features and the assignment of attribute values and may be introduced at various stages of data compilation and database development. Models of uncertainty have been proposed for spatial information that incorporate ideas from natural language processing, the value of information concept, non-monotonic logic and fuzzy sets, and evidential and probability theory [51]. In modern GIS there is a need to more precisely model and represent the underlying uncertain spatial data. Models have been proposed recently allowing enriching database models to manage uncertain spatial data. A major motivation for this is that there exist geographic objects with uncertain boundaries, and fuzzy sets are a natural way to represent this uncertainty [11]. An ontology for spatial data has been developed in which the terms imperfection, error, imprecision and vagueness are organized into a hierarchy to assist in management of these issues [19]. At the most basic level of vagueness modeling approaches for spatial data are considered including fuzzy set and rough set theory.

The following section discusses uncertainty and how rough set uncertainty can be managed in databases, as well as the rough set modeling of spatial data. Section 3 provides an overview of various types of representations of spatial phenomena using fuzzy and rough set techniques. The representation of spatial relationships is discussed in Section 4, along with the management of uncertainty in these relationships. In Section 5 data mining for uncertain data is discussed. Lastly, conclusions and directions for future research are presented.

2 Background

In this section we discuss some of the approaches to modeling uncertainty in spatial data using fuzzy and rough set theory. Then we provide a brief introduction to the basic concepts and terminology of fuzzy set and rough set theory.

2.1 Overview

In general, the idea of implementing fuzzy set theory as a way to model uncertainty in spatial databases has a long history. Some early work by geographical scientists in the 1970s utilized fuzzy sets [61] in topics such as behavioral geography and geographical decision making [23]. However, the first consistent approach to the use of fuzzy set theory as it could be applied in GIS was developed by Robinson [39]. He has considered several models as appropriate for this situation—two early fuzzy database approaches using simple membership values in relations, and a similarity-based approach. In modeling a situation in which both the data and relationships are imprecise, he assesses that this situation entails imprecision intrinsic to natural language which is possibilistic in nature. For example if we are classifying various slopes in a particular region and wish to use a fuzzy set representation of steep slopes then we might have the start of steepness as *a = 15 degrees* and *b= 30 degrees* for slopes that are certainly classified as steep, i.e., have membership value of 1. Another application is in soil classification as a certain soil sample may have 0.49 membership in the set of Loamy Soil, it may have 0.33 membership in Sandy Soil, and it may have 0.18 membership in Rocky Soil. Another spatial modeling approach considers some objects as comprising a core (full membership of 1.0 in the set in question), or a boundary (the area beyond which they have no or negligible membership in the set). A classic spatial example of the core and boundary problem is determining where a forest begins. Is it determined based on a hard threshold of trees per hectare? This may be the boundary set by management policy, but it is likely not the natural definition. There are several ways to manage these uncertain boundaries [22]. If a spatial database can represent the outlying trees as being partial members of the forest, then a decision maker will see these features as being partial members if the database is queried or the data presented on a graphical user interface.

More recently, there have been a number of efforts utilizing fuzzy sets for spatial databases including: capturing spatial relationships [12], querying spatial information [55], and object-oriented modeling [14]. Models have been proposed in recent years that allow for enriching database models to manage uncertain spatial data [35]. A major motivation for this is that there exist geographic objects with uncertain boundaries, and fuzzy sets are a natural way to represent this uncertainty.

A description of spatial data using rough sets was proposed in the ROSE system [41], which focused on a formal modeling framework for realm-based spatial data types in general. In [58] Worboys models imprecision in spatial data based on the resolution at which the data is represented, and for issues related to the integration of such data. This approach relies on the issue of indiscernability – a core concept for rough sets – but does not carry over the entire framework and is just described as "reminiscent of the theory of rough sets" [59]. Ahlqvist and colleagues [2] used a rough set approach to define a rough classification of spatial data and to represent spatial locations. They also proposed a measure for quality of a rough classification compared to a crisp classification and evaluated their technique on actual data from vegetation map layers. They considered the combination of fuzzy and rough set approaches for reclassification as required by the integration of geographic data. Another research group in a mapping and GIS context [57] have developed an approach using a rough raster space for the field representation of a spatial entity and evaluated it on a classification case study for remote sensing images. In [10] Bittner and Stell consider K-labeled partitions, which can represent maps, and then develop their relationship to rough sets to approximate map objects with vague boundaries. Additionally they investigate stratified partitions, which can be used to capture levels of details or granularity such as in consideration of scale transformations in maps, and extend this approach using the concepts of stratified rough sets.

2.2 Fuzzy Set Basics

Extensions to ordinary set theory, known as fuzzy set theory, provide widely recognized representations of imprecision and vagueness [61]. Here we will overview some basic concepts of fuzzy sets and a more complete introduction can be found in several comprehensive sources [18, 29, 38].

Conventionally we can specify a set C by its characteristic function, $\mathrm{Char}_C(x)$. If U is the universal set from which values of C are taken, then we can represent C as

$$C = \{ x \mid x \in U \text{ and } \mathrm{Char}_C(x) = 1 \}$$

This is the representation for a crisp or non-fuzzy set. For an ordinary set C, the characteristic function is of the form

$$\mathrm{Char}_C(x): U \rightarrow \{0, 1\}$$

However for a fuzzy set A we have

$$\mathrm{Char}_A(x): U \rightarrow [0, 1]$$

That is, for a fuzzy set the characteristic function takes on all values between 0 and 1 and not just the discrete values of 0 or 1 representing the binary choice for membership in a conventional crisp set such as C. For a fuzzy set the characteristic function is often called the membership function and denoted $\mu_A(x)$. As an example of a fuzzy set consider a description of mountainous terrain. We want to use a linguistic terminology to represent whether an estimate of elevation is viewed as a low, medium, or high cost. If we assume we have obtained opinions of experts knowledgeable about such terrain, we can define fuzzy sets for these terms. Clearly it is reasonable to represent these as fuzzy sets as they represent judgmental opinions and cannot validly be given precise specification. Here we will provide a typical representation of a fuzzy set for the term "HIGH".

$$\mathrm{HIGH} = \{ 0.0 / 0.1\mathrm{K}, 0.125 / 0.5\mathrm{K}, 0.5 / 1\mathrm{K}, 0.8 / 2\mathrm{K}, 0.9 / 3\mathrm{K}, 1.0 / 4\mathrm{K} \}$$

This typical representation enumerates selected elements and their respective membership values as $x / \mu_A(x)$. The elements are shown in kilometers, i.e., K. It is also common to more fully specify the membership function $\mu_A(x)$ in an analytic form or as a graphical depiction. The membership function for the representation shown as in HIGH could be fully specified by interpolation between the consecutive elements listed. Also extrapolation past the first and last elements completes the specification, i.e.,

$$\mu_A(x) = 0.0, \quad x \leq 0.1\mathrm{K} \text{ and } \mu_A(x) = 1.0, \quad x \geq 4\mathrm{K}$$

All of the basic set operations must have equivalent ones in fuzzy sets, but there are additional operations based on membership values of a fuzzy set that hence have no correspondence in crisp sets. We will use the membership functions μ_A

and μ_B to represent the fuzzy sets A and B involved in the operations to be illustrated.

Set Equality: $A = B \Leftrightarrow \mu_A(x) = \mu_B(x)$

Set Containment: $A \subseteq B \Leftrightarrow \mu_A(x) \leq \mu_B(x)$

Set Complement: $\overline{A} \Leftrightarrow \{ x / (1 - \mu_A(x)) \}$

For ordinary crisp sets $A \cap \overline{A} = \emptyset$; however, this is not generally true for a fuzzy set and its complement. This may seem to violate the law of the excluded middle, but this is just the essential nature of fuzzy sets. Since fuzzy sets have imprecise boundaries, we cannot place an element exclusively in a set or its complement. This definition of complementation has been justified more formally by Bellman and Giertz [7].

Set Union and Set Intersection

$$A \cup B \Leftrightarrow \mu_{A \cup B}(x) = \mathrm{Max}(\mu_A(x), \mu_B(x))$$
$$A \cap B \Leftrightarrow \mu_{A \cap B}(x) = \mathrm{Min}(\mu_A(x), \mu_B(x))$$

The justification for using the Max and Min functions for these operations is given in [7]. With these definitions, the standard properties for crisp sets of commutativity, associativity, and so forth, hold for fuzzy sets. There have been a number of alternative functions proposed to represent set union and intersection [18, 60]. For example, in the case of intersection, a product definition, $\mu_A(x) * \mu_B(x)$, has been considered.

2.3 Rough Set Basics

Rough set theory, introduced by Pawlak [37] is a technique for dealing with uncertainty and for identifying cause-effect relationships in databases as a form of database learning. They have been widely used in data mining applications. Rough sets involve the following:

U is the universe, which cannot be empty,
R is the indiscernability relation, or equivalence relation,
A = (U,R), an ordered pair, is called an approximation space,
$[x]_R$ denotes the equivalence class of R containing x, for any element x of U,
elementary sets in A - the equivalence classes of R,
definable set in A - any finite union of elementary sets in A.

Therefore, for any given approximation space defined on some universe U and having an equivalence relation R imposed upon it, U is partitioned into equivalence classes called elementary sets which may be used to define other sets in A. Given that $X \subseteq U$, X can be defined in terms of definable sets in A as following:

lower approximation of X in A is the set $\underline{R}X = \{x \in U \mid [x]_R \subseteq X\}$

upper approximation of X in A is the set $\overline{R}X = \{x \in U \mid [x]_R \cap X \neq \varnothing\}$.

Another way to describe the set approximations is as follows. Given the upper and lower approximations $\overline{R}X$ and $\underline{R}X$, of X a subset of U, the R-positive region of X is $POS_R(X) = \underline{R}X$, the R-negative region of X is $NEG_R(X) = U - \overline{R}X$, and the boundary or R-borderline region of X is $BNR(X) = \overline{R}X - \underline{R}X$. X is called R-definable if and only if $\underline{R}X = \overline{R}X$. Otherwise, $\underline{R}X \neq \overline{R}X$ and X is rough with respect to R. In Figure 1 the universe U is partitioned into equivalence classes denoted by the squares. Those elements in the lower approximation of X, $POS_R(X)$, are denoted with the letter P and elements in the R-negative region by the letter N. All other classes belong to the boundary region of the upper approximation.

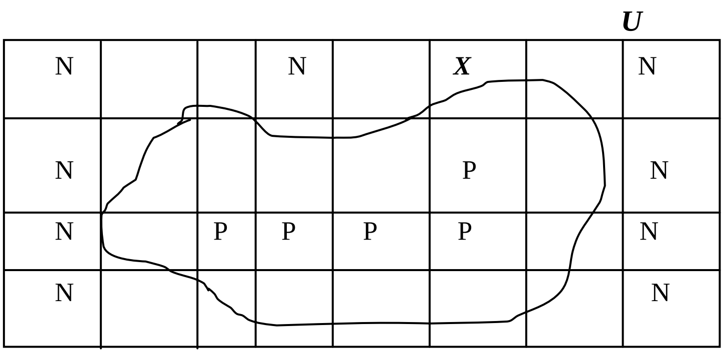

Fig. 1 Example of a Rough Set X

2.4 *Rough Set Modeling of Spatial Data*

Let U = {tower, stream, creek, river, forest, woodland, pasture, meadow}and let the equivalence relation R be defined as follows:

R* = {[tower], [stream, creek, river], [forest, woodland], [pasture, meadow]}.

Given some set X = { tower, stream, creek, river, forest, pasture}, we would like to define it in terms of its lower and upper approximations:

$\underline{R}X$ = {tower, stream, creek, river}, and

$\overline{R}X$ = {tower, stream, creek, river, forest, woodland, pasture, meadow}.

The lower approximation contains those equivalence classes that are included entirely in the set X. The upper approximation contains the lower approximation plus those classes that are only partially included in X. In this example all the values in the classes [tower] and [stream, creek, river] are included in X so these belong to the lower approximation region. The class [forest, woodland] is not entirely included in X since X does not contain 'woodland.' However, [forest, woodland] is part of the upper approximation since forest $\in$ X. A rough set in A is the group of subsets of U with the same upper and lower approximations. In the example given, the rough set is

{{tower, stream, creek, river, forest, pasture}

{tower, stream, creek, river, forest, meadow}

{tower, stream, creek, river, woodland, pasture}

{tower, stream, creek, river, woodland, meadow}}.

Although the rough set theory defines the set in its entirety this way, for our applications we typically will be dealing with only certain parts of this set at any given time. The major rough set concepts of interest are the use of an indiscernibility relation to partition domains into equivalence classes and the concept of lower and upper approximation regions to allow the distinction between certain and possible, or partial, inclusion in a rough set.

The indiscernibility relation allows us to group items based on some definition of 'equivalence' as it relates to the application domain. We may use this partitioning to increase or decrease the granularity of a domain, to group items together that are considered indiscernible for a given purpose, or to "bin" ordered domains into range groups. In order to allow possible results, in addition to the obvious, certain results encountered in querying an ordinary spatial database system, we may employ the use of the boundary region information in addition to that of the lower approximation region. The results in the lower approximation region are certain. These correspond to exact matches. The boundary region of the upper approximation contains those results that are possible, but not certain.

3 Applications

There have been many applications of both fuzzy and rough set theory to various topics related to spatial data. In following we discuss a number of these important applications and present details on significant ones.

3.1 Fuzzy Set Terrain Modeling

Several approaches to deriving fuzzy representation of terrain features from digital elevation models (DEM) have been proposed. Skidmore [47] used Euclidean distances of a given location to the nearest streamline and ridgeline to represent the location's relative position, but a Euclidean distance is often not sufficient to represent local morphological characteristics. Irvin et al. [27] performed a continuous classification of terrain features using the fuzzy k-mean method. As a basically unsupervised classification, the fuzzy k-mean method sometimes has difficulty in producing results that satisfactorily match domain experts' (e.g., soil scientists) views on landscapes. MacMillan et al. [33] developed a sophisticated and comprehensive rule-based method for fuzzy classification of terrain features that requires intensive terrain analysis operations and has a high demand for users' knowledge of local landform.

Another method [45] derives the fuzzy membership of a test location as being a specific terrain feature based on the location's similarity to the typical locations of that terrain feature. This can be very useful for special terrain features that have very unique meanings to soil-landscape analysts as unique soil conditions often exist at such locations. A definition-based and a knowledge-based approach are given as ways to specify typical locations. Where there is a clear geomorphology, simple rules based on the definitions can be used to determine the typical locations. For example there are algorithms for determining ridgelines and streamlines that can be used. However, if a terrain feature has only has a local or regional meaning, finding the typical location may require knowledge from local experts. This may be captured through manual delineation using a GIS visualization tool.

The similarities of any other location to those specified typical locations can be evaluated based on a set of selected terrain attributes such as elevation, slope gradient, curvatures, etc. The process of assigning fuzzy membership value to a location then consists of three steps:

1. Evaluation of similarity of a test location and a typical location at the individual terrain attribute level.
2. Integration of similarities on individual terrain attributes yielding overall similarity between test location and a typical location.
3. Integration of test location's similarities to all typical locations producing a final fuzzy membership of the test location for being the terrain feature under concern.

3.2 Rough Sets for Gridded Data

Often spatial data is associated with a particular grid. The positions are set up in a regular matrix-like structure and data is affiliated with point locations on the grid. This is the case for raster data and for other types of non-vector type data such as topography or sea surface temperature data. There is a tradeoff between the resolution or the scale of the grid and the amount of system resources necessary to store and process the data. Higher resolutions provide more information, but at a cost of memory space and execution time.

If we approach the data from a rough set point of view, we can see that there is indiscernibility inherent in the process of gridding or rasterizing data. In Figure 2, for example, there are grid locations that represent the various lake, chemical plant, forest, boatyard, residential, and other classifications. Some grid points are directly on one of these classifications and some are in between one or more of them. A data item at a particular grid point in essence may represent data near the point as well. This is due to the fact that often point data must be mapped to the grid using techniques such as nearest-neighbor, averaging, or statistics. We may

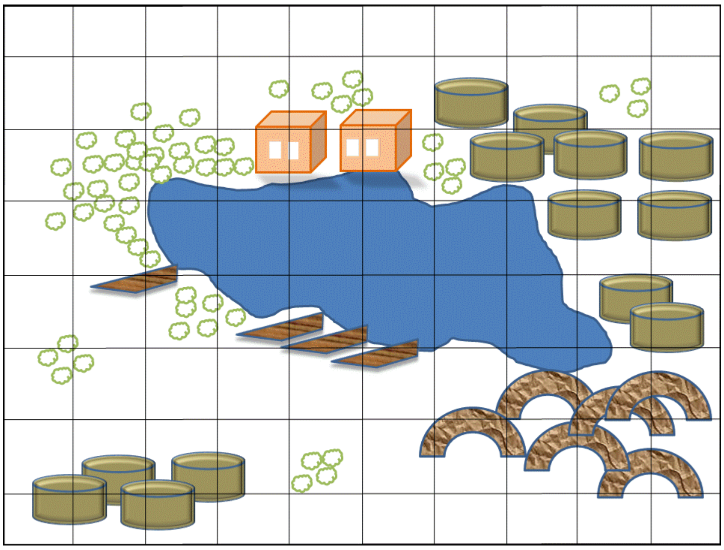

Fig. 2 Gridded data for land classification showing coarse grid lines

set up our rough set indiscernibility relation so that the entire spatial area is partitioned into equivalence classes where each point on the grid belongs to an equivalence class. If we change the resolution of the grid, we are in fact, changing the granularity of the partitioning, resulting in fewer, but larger classes.

3.3 Fuzzy Triangulated Irregular Networks

Triangulated Irregular Networks (TINs) are one common approach to represent field data as opposed to object-based spatial data. A TIN is based on a partition of the two-dimensional space into non-overlapping triangles. Extensions of TINs [54] have been developed using fuzzy membership grades, fuzzy numbers and type-2 fuzzy sets. The ETIN structure uses a mapping function that specifies a property *F* of a geographic area. Consider the description of a specific site under evaluation for purchase as “Close to” New York. So a value of 1 for the function indicates the site is *near* (or in) New York; 0 means the location is actually far (not close) from New York and intermediate values such as 0.6 implies the site might be considered as being more or less close to the city.

Another TIN extension is based on fuzzy numbers with triangular membership functions, as these provide a simple model for a fuzzy number. To use fuzzy numbers in the ETIN, it is necessary to extend the type with the associated data value in a point from a simple (crisp) value to a fuzzy set. This can be accomplished at every point of the region under consideration by associating a triangular membership function. Three characterizing points are then of importance: the two points where the membership grade equals 0 which delimit the membership function, and the intermediate point for which the membership grade equals 1.

Finally the ETIN structure can use type-2 fuzzy sets, a generalization of regular fuzzy sets, allowing imprecision as well as uncertainty regarding the membership grades to be modeled. Consider the certainty about the extent to which a site is "close to" New York. When describing for example the location of some individual, there might be doubt as to exactly where they are located. The person could be located close to New York, but also near Newark, New Jersey. A type-2 fuzzy set allows this doubt to be modeled: With this approach, the membership grade on every location is extended to a "fuzzy" membership grade. As a result, every point will now have an associated fuzzy set over [0,1].

3.4 Fuzzy Spatial Interpolation

Since as we have seen geographical data are a combination of fuzzy and crisp data types there is a need to rely on the application of fuzzy based interpolation techniques. When interpolation data are not sets of real numbers but ranges of values whose distribution within the range are qualitative, sample data have to be determined with a theory of possibility. For example, geological data may be collected from wells where it is not obvious from the sample description the exact component percentages of clay, sand, or silt. A fuzzy interpolation approach [17] is derived from gradual rules that in fact fully capture the interpolation process. The formulations are given on the basis of linear interpolation that uses fuzzy and precisely known – crisp data which has roots in the fuzzy Lagrange interpolation theorem. This approach has been applied to two dimensional spatial interpolation based on fuzzy Voronoi diagrams, fuzzy function estimator, three dimensional spatial interpolation based on fuzzy neural networks, and GIS based fuzzy spatio-temporal interpolation.

For example a fuzzy Voronoi approach can be applied to thematic maps represented by polygons with categories such as forest types where each polygon is assigned specific attributes (e.g. wood volume). Polygon boundaries are uncertain because of varying interpretations of imagery data. Distributions of attribute values over surfaces are not reliable because of sparseness of in situ measurements. Since most geographic attributes are not of a continuous nature, spatial interpolation is needed to create a continuous surface of selected attributes and to represent the transition zones between polygons. These issues can be resolved using fuzzy Voronoi diagrams first by constructing Voronoi diagrams around known points with well-specified attributes. The next step positions a “query point” in the Voronoi diagram and a new diagram reconstructed as if the query point was one of the original data points. Thus, new polygons are delineated containing the area stolen from the original polygons. The percentage of the stolen area from each polygon constitutes the fuzzy membership value for a thematic category represented by the corresponding original polygon. If a grid of query points is processed over the entire surface at regular intervals, a series of grid points with fuzzy membership values are produced for each geographic category. Linear interpolation can then be used to produce a continuous surface that can be stored in a raster GIS format. The attributes of interest are evaluated at any location on the defined fuzzy map by multiplying the mean estimated volume of the particular attribute for each geographic category by the corresponding fuzzy membership value over all geographic categories.

In a GIS spatial data is represented by snapshot layers corresponding to time intervals limiting by the temporal granularity spatial change detection. Fuzzy temporal interpolation [16] uses fuzzy probable trajectories of gradual progression from one class to another. The degrees of membership in a specific class at a particular intermediate space-time location are calculated using fuzzy set membership functions.

In [24] a spatial interpolation technique is described that is based on conservative fuzzy interpolation reasoning for interpolating fuzzy rules in sparse fuzzy rule bases. The technique works best in local spatial interpolation so a self-organizing map is used to divide the data into subpopulations in order to reduce the complexity of the whole data space to more homogeneous local regions.

4 Representation of Spatial Relations

Relationships among spatial objects can generally be classified in three types:

1. Topological - Touches, Disjoint, Overlap, …
   French border touches German border
2. Directional - East, North-West, …
   Prague is East of Frankfurt
3. Metric – Distance
   Wien is about 50 kilometers from Bratislava

Many topological relations between two objects A and B can be specified using the 9-intersection model which uses the intersections between the interior, boundary and exterior of A and B [21]. This section will describe a variety of approaches introducing uncertainty into these relationships.

4.1 Spatial Relations

In [36], Papadias and his colleagues present an approach for determining configuration similarity for spatial constraints involving topology, direction and distance. The approach utilizes extended objects for direction and topology, and centroids for distance. They handle uncertainty in the areas of fuzzy relations, e.g., an object that satisfies more than one directional constraint, as well as fuzziness related to linguistic relationship terms. The concept of graded sections, allows comparison of alternative conceptualizations of direction [30]. To describe graded sections, section bundles are introduced, providing a formal means to (1) compare alternative candidates related via a direction relation like “north” or “south-east,” (2) distinguish between good and not so good candidates, and (3) select a best candidate. Vazirgiannis [53] also considers the problem of representing uncertain topological, directional, and distance relationships on the assumption of crisply bounded objects. All relationship definitions for this approach are centroid-based. A minimal set of topological relations, overlapping and adjacency, are defined based on

Egenhofer's boundary/interior model. This model is enhanced by providing degrees of relationship satisfaction. Direction relations are defined by a sinusoidal function based on the angle between two objects' centroids. *Close* and *far* are the two categorizations of distance directions. Membership assignment to one of these categories is determined by the ratio of the distance to a maximum application-dependent distance. The three relationships are combined for query retrieval. Afterward, a similarity measure is computed for each relationship and then combined into a single, overall similarity measure. Another approach to spatial relations uses the histogram of forces [34] to provide a fuzzy qualitative representation of the relative position between two dimensional objects. This can also be used in scene description where relative positions are represented by fuzzy linguistic expressions. In Guesgen [25] we see the introduction of several approaches for reasoning about fuzzy spatial relations, including an extension of Allen's algorithm and additionally methods for fuzzy constraint satisfaction. Also relevant is [20] which presents a unified framework for approximate spatial and temporal reasoning using topological constraints as the representation schema and fuzzy logic for representing imprecision and uncertainty. The application of the resulting fuzzy representation to each of Allen's interval relationships is developed as the possibility of the occurrence of the conditions of the original definition.

Yet another approach of Cobb and Petry [12] is based on minimum bounding rectangles (MBRs) and Allen's relationships. An MBR is an approximation of the geometry of spatial objects and is defined as the smallest X-Y parallel rectangle which completely encloses an object. The use of MBRs in geographic databases is widely practiced as an efficient way of locating and accessing objects in space. An extension into the spatial domain of Allen's temporal relationships [1] represents any relationship that can exist between two one-dimensional (temporal) intervals including: before, equal, meets, overlaps, during, starts, and finishes, along with their inverses.

Given the minimum bounding rectangles of two objects, the binary relationship between the objects in both the horizontal and vertical directions can be completely defined by a tuple, [rx, ry], where rx is the one of the described above Allen's temporal relations that defines the interaction of the object MBRs in the x direction, and ry represents the same for the y direction. For example, for the case of the relationship, A [finishes, starts] B, the definition is given as:

$$\{ B_{xl} < A_{xl} < B_{x2}, A_{x2} = B_{x2}, B_{y}l < A_{y2} < B_{y2}, A_{yl} = B_{y1} \}$$

where {xl,y1} and {x2, y2} represent the lower left and upper right corners, respectively, of the minimum bounding rectangles. In Figure 3 is an example set of four object MBRs, {A,B,C,D}. A subset of the existing relationships between them consists of:

{A [before, overlaps] B; B [before, overlaps $^{-1}$] C; D [during, meets] C }.

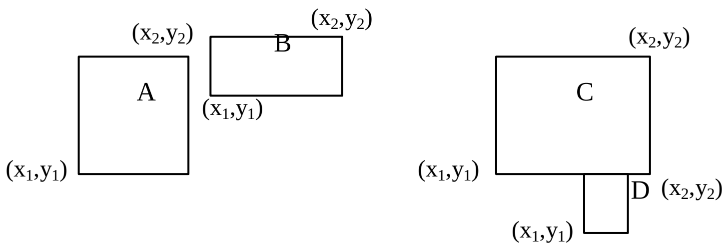

Fig. 3 Object for MBR Relationship Description

Again we can use the notation of representing one of Allen's relations by its initial letter and so we have determined for the relation *partially-surrounded-by* :

$$\{ [df]\ [fd]\ [do]\ [ds]\ [od]\ [sd]\ [do^{-1}] \}$$

These basic relationship definitions can be used in a similar manner for defining directional relationships: N, S, E, W, NE, SE, SW, NW. Given the spatial extent of two-dimensional objects, it is very likely that in any one case, more than one of the eight directions listed above will apply, to either a greater or lesser degree. So a method for defining directional relationships that would allow for fuzzy querying of any of the directional relationships that exists between two objects is needed The concept of object sub-groups is then used as a basis for determining the set of directions that defines the directional relationship between two objects.

Definitions for directions can now be defined in a manner analogous to the way in which qualitative topological relationships were defined earlier. The definition for any particular direction includes the set of all relationships containing that direction as a member of its direction set. The definition for the direction East is shown below as an example.

$$E ::=\{[dd],[df],[fd],[do],[ds], [ff],[d=],[fo],[fs],[f=],[dd^{-1}],[do^{-1}], \ldots\}$$

The basic relationship definitions and their use in defining relevant directional and qualitative topological relationships can then be used to provide a framework for the *abstract spatial graph* (ASG), a spatial data structure specifically designed to retain orientation and topological information with respect to two-dimensional objects, and to provide information to support fuzzy querying capabilities on these relationships.

The ASGs categorize the original relationships according to the level of interaction of the MBRs into four distinct categories: disjoint, tangent, overlapping and containment.

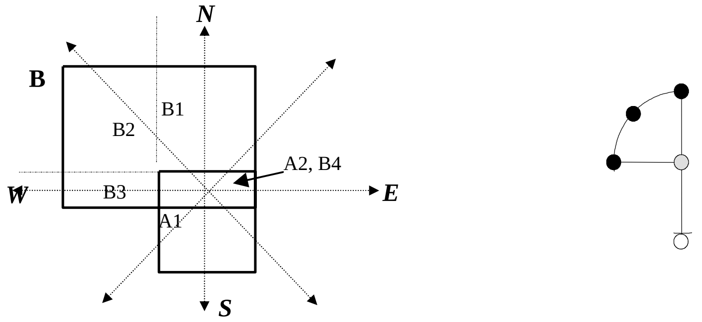

Fig. 4 Application of thresholding for ASG construction of [fo] relationship

Figure 4 shows the construction of an abstract spatial graph for the [fo] relationship using a thresholding technique. We will note that the northeastern and northwestern axes for sub-group B1, as well as the southeastern and southwestern axes for sub-group A1, are discarded, so that the node on the northern axis of the ASG singularly represents B1; the node on the northwestern axis represents B2; the node on the western axis represents B3; and the node on the southern axis represents A1. The center node of the ASG represents the sub-groups A2 and B4—the reference area. In addition to providing information directly relevant to the representation of the abstract spatial graph, we also needed to represent ancillary information that can be used for fuzzy query inferences. This information is represented in the form of node "weights" that can then be used for the defining of both fuzzy topological and directional qualifiers for use with a fuzzy query framework.

Calculation of weights uses both the areas of object sub-groups and the lengths of axes that pass through object sub-groups. Three different types of weights are computed: *axis weights*, *area weights* and *node weights*. The area weights and total node weights of ASGs directly support fuzzy queries regarding qualitative topological and directional information in two specific ways. Area weights provide an indication of the degree to which an object participates in a qualitative topological relationship. By mapping ranges of area weights to linguistic qualifiers such as *some*, *most*, etc., fuzzy information such as "some of object A overlaps most of object B," can be determined.

Total node weights, on the other hand, are used to indicate the extent to which one object can be considered to lie in a certain direction with respect to a second object. Again, ranges of weights can be correlated to linguistic terms, e.g. *slightly*, *mostly*, to provide qualifiers for directional orientation. Then, for example, one could determine that, while object A is slightly southwest of object B, it is at the same time mostly west of object B.

So we can determine for our example of Figure 3 that:

1. B is *mostly* west of C
2. *Little* of B is northeast of A
3. D is *directly* south of C
4. C is *slightly* southeast of B

4.2 Topological Spatial Relationships for Vague Regions

The interplay of topological relations and nearness lies at the core of the motivation of the formalism developed in a series of papers by Schockaert et.al [42 43,44]. These papers provide characterizations of the fuzzy spatial relations, corresponding to the particular case where connection is defined in terms of closeness between fuzzy sets. Also generalization of region connection calculus (RCC) is based on fuzzy set theory, and a development of how reasoning tasks such as satisfiability and entailment checking can be cast into linear programming problems.

Keukelaar [28] develops an approach for rough spatial topological relations using 3-valued logic, allowing "maybe" answers to queries about the spatial relationships between objects. Wang et.al [56] deal with imprecise spatial relationships in a straightforward manner for the 9-intersection model. In this they replace the interior, exterior and boundary with positive, negative and boundary regions in a rough set sense based on the lower and upper approximations. A rough matrix representation facilitates computation of rough topological relationships among several spatial objects. In Zhan [62] a method is developed for approximately analyzing binary topological relations between geographic regions with indeterminate boundaries. It shows the eight binary topological relations between regions in a two-dimensional space can be easily determined by this method.

A computational fuzzy topology can be developed based on the interior operator and closure operator [32]. These operators are further defined as a coherent fuzzy topology—the complement of the open set is the closed set and vice versa; where the open set and closed set are defined by interior and closure operators—two level cuts. The elementary components of fuzzy topology for spatial objects—interior, boundary and exterior—are thus computed based on the computational fuzzy topology. Yet another approach proposes basic fuzzy spatial object types based on fuzzy topology [52]. These object types are a natural extension of current non-fuzzy spatial object types. A fuzzy cell complex structure is defined for modeling fuzzy regions, lines and points. Furthermore, fuzzy topological relations between these fuzzy spatial objects are formalized based on the 9-intersection approach.

In [9] Bittner and Stell present an approach to spatial relations where the consideration of uncertainty is based on the case in which there is limited resolution of spatial data and using approximations that have a close relationship to rough sets. They develop two methods for approximating topological relations, syntactic and semantic. In the first, use is made of the set of precise regions which could be an interpretation of the approximate regions. The syntactic approach also uses algebraic operations which generalize operations on precise regions by using pairs of greatest minimal and least maximal meet operations to approximate the crisp meet used for defining topological relations.

Rough set [4] and egg-yolk [13] approaches can also be used to model spatial relationships. In spatial data, it is often the case that we need information concerning the relative distances of objects. Is object A *adjacent to* object B? Or, is object A *near* object B? The first question appears to be fairly straightforward. The system must simply check all the edges of both objects to see if any parts of them are coincident. This provides what would be *certain* results in the ideal case . However, often in a GIS, data is input either automatically via scanners or digitized by humans, and in both cases it is easy for error in position of data objects to occur. Therefore, it might be desirable to have the system check to see if object B is very near object A, to derive a *possible* result. If so, the user could be informed that “it is not certain, but it is possible, that A is adjacent to B.” One may want to know whether a cliff is next to the sea. If the system returns the result that it is possible, but not certain, that the cliff is adjacent to the sea, for example, he may be led to investigate the influence of the tides in the area to determine whether low beaches alongside the cliffs are exposed at low tide.

The concepts of connection and overlap can be managed by rough sets in a similar manner to the above. Connection is similar to adjacency, but related to line type objects instead of area objects. Overlap can be defined in a manner similar to that of nearness with the user deciding how much overlap is required for the lower approximation. Coincidence of a single point may constitute *possible* overlap, as can very close proximity of two objects, if there is a high degree of positional error involved in the data.

Inclusion is related to overlap as follows. If an object A is completely surrounded by some other object B, perhaps we can conclude with certainty that A is included in B, lacking any additional information about the two objects. If the two objects overlap, then it may be possible that one of the objects includes the other. Approximation regions can be defined to reflect these concepts as well.

Both the rough set and egg-yolk approaches are useful for managing the types of uncertainty and vagueness related to topology, a few of which were just briefly discussed. These include concepts such as nearness, contiguity, connection, orientation, inclusion, and overlap of spatial entities.

If we are only concerned about the vagueness of boundaries, we may be inclined to use the egg-yolk approach [13], since this approach does not include any partitioning of the space into equivalence classes as does rough sets. In this approach concentric subregions make up a vague region, with inner subregions having the property that they are ‘crisper’ than outer subregions. These regions indicate a type of membership in the vague region. The simplest case, is that of two

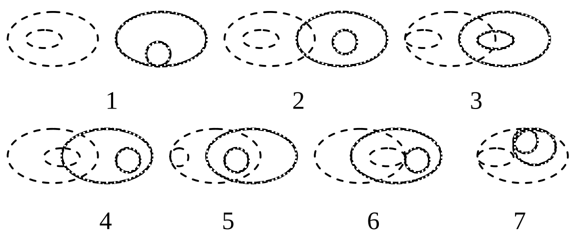

Fig. 5 A sample of the 46 possible relationships between regions X (dashed line) and Y (dotted line). A solid line indicates coincidence of an X and Y region boundary.

subregions. In this most common case, the center region is known as the yolk, the outer region surrounding the yolk is known as the white, and the entire region, as the egg. Figure 5 depicts a sample of these relationships.

The yolk and egg regions correspond to the lower and upper approximation regions of rough sets respectively. The rough set theory has only these two approximation regions, unlike the possible numerous subregions that may make up a vague region in the egg-yolk method. However, because of the indiscernibility relation in rough sets, one can vary the partitioning in order to increase or decrease the level of uncertainty present, which results in changes to the approximation regions.

Consider specifically the results of Cohn and Gotts [13] who delineate forty-six possible egg-yolk pairs showing all of the possible relationships between two vague regions. The forty-six configurations of egg-yolk pairs were clustered into thirteen groups based on RCC-5 [40] relations between complete crispings, or relations that are "mutually crispable". Each cluster relates to one or more additional clusters via a crisping relationship or a subset relationship between a set of complete crispings.

The clustering of egg-yolk pairs can also be viewed by noting that the relationships for each cluster based on mathematical principles from rough sets. We now recall that "crisping" from the egg-yolk theory can also be related to forcing a finer partitioning on the domain for the rough sets. Some definitions from rough set theory used in categorizing the clusters include:

Equality of 2 rough sets:

Two rough sets X and Y are equal, X = Y, if $\underline{R}X = \underline{R}Y$ and $\overline{R}X = \overline{R}Y$.

Intersection of two rough sets:

$\underline{R}(X \cap Y) = \underline{R}X \cap \underline{R}Y$, and $\overline{R}(X \cap Y) = \overline{R}X \cap \overline{R}Y$.

Subset relationship:

$X \subset Y$ implies that $\underline{R}X \subset \underline{R}Y$ and $\overline{R}X \subset \overline{R}Y$.

In [4] properties of rough sets are used to define the crispings in the various topological clusters as well as the spatial relationships themselves. Figure 6 shows the relationships between clusters based on the levels of crisping from one cluster to another. Numbers within each cluster represent each of the 46 egg-yolk pairs of Cohn and Gotts [13] denoting uncertain spatial relationship for two vague regions. Within the hierarchy an arrow from one cluster to another means that there is some property of rough sets theory that is added to those properties of the beginning cluster in order to make it more "crisp."

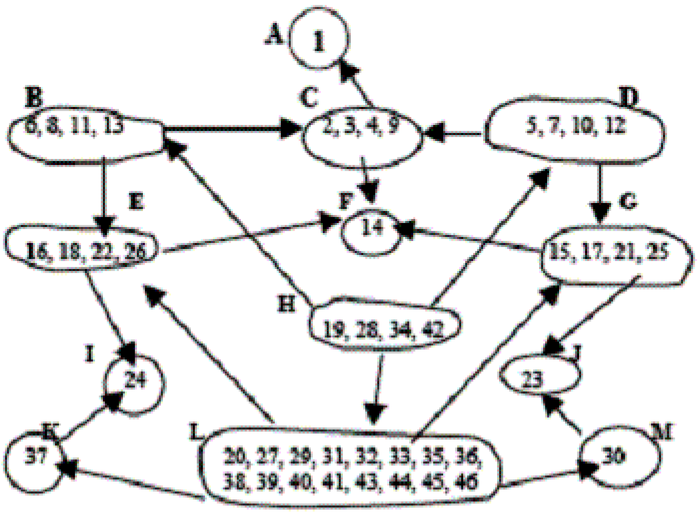

Fig. 6 Clustering of egg-yolk relationships

Spatial and geographical information systems will continue to play an ever-increasing role in applications based on spatial data. Uncertainty management will be necessary for any of these applications, and both rough sets and egg-yolk methods are appropriate for the representation of vague regions in spatial data. Rough sets, however, can also model indiscernibility and allow for the change of granularity of the partitioning through its indiscernibility relation, which has an effect on the boundaries of the vague regions, and also allows the extension of egg-yolk regions from continuous to discrete space. The clustering of egg-yolk pairs by RCC-5 relations can be expressed in terms of operations using rough sets, and rough set techniques can further enhance the egg-yolk approach. The interrelationships between rough set, egg-yolk, and RCC models merit further study.

5 Mining Spatial Information

5.1 Spatial Data Mining

An approach [31] to the discovery of association rules for fuzzy spatial data combined and extended techniques developed in both spatial and fuzzy data mining in order to deal with the uncertainty found in typical spatial data. It attempts to uncover correlations of spatially related data such as soil types, directional or geometric relationships, etc. For example an association rule that can be discovered by mining appropriate spatial data is:

> *If C is a small <u>city</u> and has good <u>terrain</u> nearby then there is a <u>road</u> nearby with 90% confidence.*

Such a rule incorporates fuzzy information in the linguistic terms used such as “small” and “nearby”.

In the spatial data mining area there have only been a few efforts using rough sets. In the research described in [3], Beaubouef et.al. have investigated approaches for attribute induction knowledge discovery in rough spatial data. Bittner [8] considers rough sets for spatio-temporal data and how to discover characteristic configurations of spatial objects focusing on the use of topological relationships for characterizations. In a survey of uncertainty-based spatial data mining, Shi et al. [46] provide a brief general comparison of fuzzy and rough set approaches for spatial data mining.

5.2 Fuzzy Minimum Bounding Rectangles

To utilize minimum bounding rectangles for vague regions, in [50] a fuzzy MBR (FMBR) is defined as consisting of nested rectangles. The inner rectangle is the MBR over the core of the vague region (certain region or membership =1). The outer rectangle is an MBR over the outer boundary of the vague region. This approach allows the consideration of common indexing approaches such as grid files or R-trees.

A vague region is one whose boundaries are or can not be precisely defined and we can consider them as being of two main components: the core and the boundary. The core and the boundary are approximated by their minimum bounding rectangle (MBR) respectively. A fuzzy representation, called Fuzzy Minimum Bounding Rectangles (FMBR) [49], can represent the different degrees of membership of the point located inside the vague region.

Geographic features are a direct representation of geographic entities rather than geometric elements such as a point, line or polygon. A feature is then defined as an entity with common attributes and relationships. The FMBR [48, 49] represents the generalization of the underlying irregular polygon delimiting the fuzzy region since the FMBR encloses all the points of the map space where our feature of interest is located.

The FMBR can be also considered as the circumscribed rectangle (CR) of the underlying fuzzy polygon. Iterative generation of inner bounding rectangles is performed until we have the inscribed rectangle (IR) of the underlying object. So, the IR is the maximum inner rectangle inside the object, and it corresponds to the core of the fuzzy region. Distances between the IR and the FMBR are used to represent the fuzzy boundary.

A spatial membership function based on Euclidean distance will be used to determine the degree of belonging of a feature to the fuzzy set. Thus, features inside the IR or core will have degree of membership of 1. This degree will be gradually decreased while we move away from the core. Points located outside of the FMBR will have a membership degree of 0.

An FMBR is a natural representation for many commonly occurring spatial situations. The problems of identifying a spatial boundary have been under considerable attention for the GIS area [11]. For example consider photo-interpreters who are trying to label a forest in an image. There is clearly a region (core) which all agree is the heart of the forest and merits the specific labeling. However, as the forest thins out into meadows all around, there is no sharp boundary delimiting the forest area. Rather the density of the trees decreases gradually until there is just open meadow land. It is just such a situation that we are trying to model by means of an FMBR.

A graphical representation of the fuzzy minimum bounding rectangle, as described above, is illustrated in Figure 7. The underlying vague region Ậ is approximated by the FMBR (Ậ). This first approximation is also called the circumscribed rectangle (CR) of the fuzzy region. In other words, the FMBR or CR corresponds to the minimal rectangle with edges parallel to the x and y axes that optimally enclose the vague region Ậ.

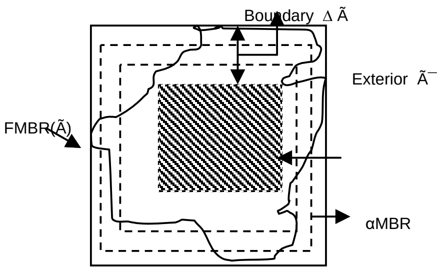

Fig. 7 FMBR Representation

αMBR-cuts allow us to make finer distinctions inside the fuzzy region since αMBR-cuts are individual crisp regions inside the FMBR Thus, we can think of a fuzzy structured region as an aggregation of crisp α-level regions. αMBRs start to be defined from the edge of the FMBR(Ậ) to the core of (Ậ). The more external the αMBR-cut the lower the degree of membership in the fuzzy set representing (Ậ) as locations which are closer to the core will have higher membership degrees. The shadowed rectangle labeled as Core corresponds to the inscribed rectangle. Since the IR is totally inside (Ậ) we assume that the points in the core belong to the fuzzy region with a membership 1.0. Details about the representation and spatial relationships of FMBRs can be found in [49], [50].

Now we can discuss an approach to an indexing structure that could be used to represent FMBRs. One commonly used index structure in spatial data bases is the R-tree [26] which is the basis of all R-tree variants. Each node corresponds to a disk page and a n-dimensional rectangle. Any entry in the tree is a pair (*ref*, *rect*), where *ref* is the address of the child node and *rect* is the MBR of all entries in that child node. The root has at least 2 children if not a leaf node. The number of entries in each node is between m (fill-factor) and M (number of entries that can fit in a node), where $2 \leq m \leq M/2$. All leaves are at the same level. Leaves contain entries of the same format, where *ref* points to a database object, and *rect* is the MBR of that object. An object appears in one, and only one of the tree leaves. R-trees are dynamic structures since insertion and deletion can be intermixed with queries and no periodic global reorganization is required. The external memory structure is multi-way and it is indexed by MBRs.

R-trees present several weaknesses mainly due to the overlap between buckets regions at the same tree level. Moreover, the region perimeters should be minimized in order to avoid insertion problems. Insertion requires multiple paths of the tree, since the inserted spatial feature may intersect more than one intermediate node, and its clipping parts should be inserted in leaves under all such nodes. R*-trees are variations that avoid some of these problems. Representing FMBRs using an R*-tree structure was found very suitable since we can take advantage of the MBR representation of the objects in this model. Figure 8 corresponds with our FMBR R*-tree description.

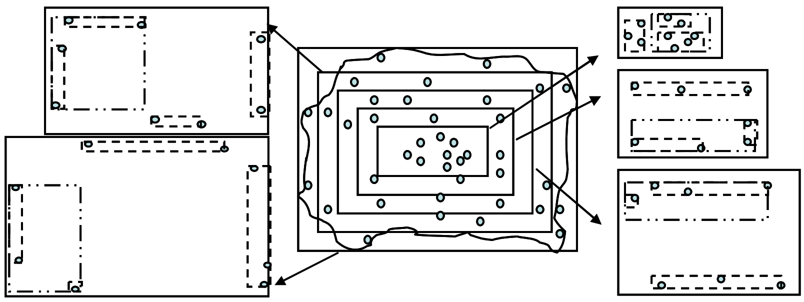

Fig. 8 Spatial Representation of αMBR-cuts

Since we are interested in treating each αMBR-cut independently we have located each of them as root nodes of the tree. This structure allows us to access the features inside the vague region with a specific degree of membership following a unique path from the root. In addition, geographically close features belonging to the same αMBR-cut can be grouped in MBRs to improve the retrieval process.

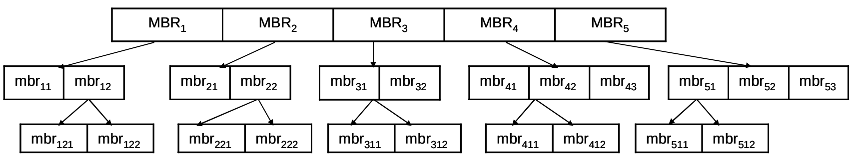

Fig. 9 FMBR R*-Tree Representation

The R-Tree of the Figure 9 contains five nodes at the root corresponding to the core, and the four αMBRs approximating the boundary. The core αMBR_1 has two MBRS: mbr_{11} and mbr_{12}, and mbr_{12} contains mbr_{121} and mbr $_{122}$. A similar structure is maintained in the remaining nodes.*

5.3 Rough Object Oriented Spatial Database

Object-oriented databases have become quite popular for many reasons. Classes and inheritance allow for code reuse through specialization and generalization, and encapsulation packages the data and methods that act on the data together in an object. Objects can be defined to represent very complex data structures and to model relationships in the data, as is often the case with spatial data. Object modeling helps in understanding the requirements of an enterprise, and object-oriented techniques lead to high quality systems that are easy to modify and to maintain. Because many newer applications involving CAD/CAM, multimedia and GIS are not suitable for the standard relational database model, object-oriented databases may be developed to meet the needs of these more complex applications.

A formal generalized model for object-oriented databases was extended to incorporate rough set techniques in [5] where the rough set concepts of indiscernibility and approximation regions were integrated into a rough object-oriented framework. In this model there is a type system, *ts*, containing literal types $T_{literal}$, which can be base types, collection literal types, or structured literal types. It also contains T_{object}, which specifies object types, and $T_{reference}$, the set of specifications for reference types.

Each domain is a subset of the set of domains, $dom_{ts} \in D_{ts}$. This domain set, along with a set of operators O_{ts} and a set of axioms A_{ts}, capture the semantics of the type specification. The type system is then defined based on these type specifications, the set of all programs P, and the implementation function mapping each type specification for a domain onto a subset of the powerset of P that contains all the implementations for the type system. Of particularly interested are object types defined as :

Class $id(id_1{:}s_1; \ldots; id_n{:}s_n)$ or

Class id: $\overline{\overline{id}}_1, \ldots, \overline{\overline{id}}_n(id_1{:}s_1; \ldots; id_n{:}s_n)$

where id, an identifier, names an object type, $\{\overline{\overline{id}}_i \mid 1 \le i \le m\}$ is a finite set of identifiers denoting parent types of t, and $\{ id_i{:}s_i \mid 1 \le i \le n\}$ is the finite set of characteristics specified for object type t within its syntax. This set includes all the attributes, relationships and method signatures for the object type. The identifier for a characteristic is id_i and the specification is s_i for each of the $id_i{:}s_i$.

Consider a GIS which stores spatial data concerning water and land forms, structures, and other geographic information. If simple types are previously defined for string, set, geo, integer, etc., then one may specify an object type as

```
Class ManMadeFeature (
        Location: geo;
        Name: string;
        Height: integer;
        Material: Set(string));
```

An example instance of the object type *ManMadeFeature* might be

> [oid1, Ø, *ManMadeFeature*, Struct(0289445, “KXYZ radio tower”, 60, Set(steel, plastic, aluminum))]

following the definition of instance of an object type [15], the quadruple o = [*oid*, *N*, *t*, *v*] consisting of a unique object identifier, a possibly empty set of object names, the name of the object type, and for all attributes, the values ($v_i \in dom_{si}$) for that attribute, which represent the state of the object. The object type t is an instance of the type system *ts* and is formally defined in terms of the type system and its implementation function t = [$ts, f_{impl}^{type}(ts)$].

Rough set uncertainty is modeled through the indiscernibility relations specified for domains and class methods for approximation region results. Each domain class *i* in the database, $dom_i \in D_i$, has methods for maintaining the current level of granulation, changing the partitioning, adding new domain values to the hierarchy, and for determining equivalence based on the current indiscernibility relation imposed on the domain class. Every domain class, then, must be able to not only store the legal values for that domain, but to maintain the grouping of these values into equivalence classes. This can be achieved through the type implementation function and class methods, and can be specified through the use of generalized constraints as in [15] for a generalized object-oriented database.

The semantics of rough set operations discussed for relational databases in [6] apply similarly for the object database paradigm. However, the implementation of these operations is done via methods associated with the individual object classes. The incorporation of rough set techniques into an object database model allow not only for the management of uncertainty in spatial data, but also for the representation of complex data relationships and the defining of methods for special cases that often exist in GIS.

6 Conclusions and Future Directions

Fuzzy and rough set approaches are increasingly being applied to many areas of spatial data. In this chapter we presented ways in which rough and fuzzy set uncertainty management may be integrated into applications involving spatial data. We reviewed rough sets, an important mathematical theory, applicable to many diverse fields. Rough sets have predominantly been applied to the area of knowledge discovery in databases, offering a type of uncertainty management different from other methods such as probability, fuzzy sets, and others. Both rough set and fuzzy set theory can also be applied to database models.

The chapter also discussed the use of rough and fuzzy set techniques for the representation of spatial data relationships, terrain modeling, gridded data, triangulated irregular networks, and spatial interpolation. Their use in the modeling of topological spatial relationships for vague regions was presented, and their integration into and data mining of object-oriented and other spatial databases discussed. The main emphasis for future work is the incorporation of some of these research topics into mainstream GIS commercial products.

Acknowledgments The authors would like to thank the Naval Research Laboratory's Base Program, Program Element No. 0602435N for sponsoring this research.

References

[1] Allen, J.: Maintaining Knowledge about Temporal Intervals. Comm. of the ACM 26(11), 832–843 (1983)
[2] Ahlqvist, O., Keukelaar, J., Oukbir, K.: Rough classification and accuracy assessment. International Journal of Geographical Information Science 14(5), 475–496 (2000)
[3] Beaubouef, T., Ladner, R., Petry, F.: Rough Set Spatial Data Modeling for Data Mining. Int. Jour. of Intelligent Systems 19, 567–584 (2004)
[4] Beaubouef, T., Petry, F., Ladner, R.: Spatial Data Methods and Vague Regions: A Rough Set Approach. Applied Soft Computing Journal 7, 425–440 (2007)
[5] Beaubouef, T., Petry, F.: Uncertainty in an OODB Modeled by Rough Sets. In: The 9th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2002), Annecy, France, July 2002, pp. 1697–1703 (2002)
[6] Beaubouef, T., Petry, F., Buckles, B.: Extension of the Relational Database and its Alge-bra with Rough Set Techniques. Computational Intelligence 11(2), 233–245 (1995)
[7] Bellman, R., Giertz, M.: On the Analytic Formalism of the Theory of Fuzzy Sets. Information Sciences 5, 149–156 (1973)
[8] Bittner, T.: Rough sets in spatio-temporal data mining. In: Roddick, J., Hornsby, K.S. (eds.) TSDM 2000. LNCS (LNAI), vol. 2007, pp. 89–104. Springer, Heidelberg (2001)
[9] Bittner, T., Stell, J.: Approximate Qualitative Spatial Reasoning. Spatial Cognition and Computation 2(4), 435–466 (2002)
[10] Bittner, T., Stell, J.: Stratified Rough Sets and Vagueness. In: Kuhn, W., Worboys, M.F., Timpf, S. (eds.) COSIT 2003. LNCS, vol. 2825, pp. 286–303. Springer, Heidelberg (2003)
[11] Burrough, P.: Natural objects with indeterminate boundaries. In: Burrough, P.A., Frank, A. (eds.) Geographic Objects with Indeterminate Boundaries, pp. 3–28. Taylor & Francis, London (1996)
[12] Cobb, M., Petry, F.: Modeling Spatial Data within a Fuzzy Framework. Jour. of Amer. Soc. Information Science 49(3), 253–266 (1998)
[13] Cohn, A., Gotts, N.: The 'Egg-Yolk' Representation of Regions with Indeterminate Boundaries. In: Burrough, P., Frank, A. (eds.) Geographic Objects with Indeterminate Boundaries GISDATA II. European Science Foundation, ch. 12 (1996)
[14] Cross, V., Firat, A.: Fuzzy objects for geographical information systems. Fuzzy Sets and Systems 113, 19–36 (2000)
[15] De Tré, G., De Caluwe, R.: A Generalized Object-Oriented Database Model with Gen-eralized Constraints. In: Proc. of NAFIPS 1999, New York, pp. 381–386 (1999)
[16] Dragicevic, S., Marceau, D.: A fuzzy set approach for modelling time in GIS. International Journal of Geographical Information Science 14(3), 225–245 (2000)

[17] Dragićević, S.: Multi-Dimensional Interpolations with Fuzzy Sets. In: Petry, F., Robinson, V., Cobb, M. (eds.) Fuzzy Modeling with Spatial Information for Geographic Problems, pp. 143–158. Springer, Heidelberg (2005)
[18] Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications. Academic Press, NY (1980)
[19] Duckham, M., Mason, K., Stell, J., Worboys, M.: A formal ontological approach to imperfection in geographic information. Computers, Environment and Urban Systems 25, 89–103 (2001)
[20] Dutta, S.: Approximate spatial reasoning: Integrating qualitative and quantitative constraints. International Journal of Approximate Reasoning 5(3), 307–331 (1991)
[21] Egenhofer, M., Franzosa, R.: Point set Topological Spatial Relations. Int. Journal of Geographical Information Systems 5(2), 161–174 (1991)
[22] Fisher, P.: Boolean and fuzzy regions. In: Burrough, P.A., Frank, A. (eds.) Geographic Objects with Indeterminate Boundaries, pp. 87–94. Taylor and Francis, London (1996)
[23] Gale, S.: Inexactness, fuzzy sets, and the foundations of behavioral geography. Geographical Analysis 4, 337–349 (1972)
[24] Gedeon, T., Wong, K., Wong, P., Huang, Y.: Spatial Interpolation using Fuzzy GIS Ency Reasoning. Transactions in GIS 7(1), 55–66 (2003)
[25] Guesgen, H.: Fuzzifying Spatial Relations. In: Matsakis, P., Sztandra, L. (eds.) Applying Soft Computing in Defining Spatial Relations, pp. GR 99-122. Physica Verlag, Heidelberg (2002)
[26] Guttman, A.: R-trees: A Dynamic Index Structure for Spatial Searching. In: Proc. ACM SIGMOD International Conference on Management of Data, pp. 47–57 (1984)
[27] Irvin, B., Ventura, S., Slater, B.: Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin. Geoderma 77, 137–154 (1996)
[28] Keukelaar, J.: Topics in Soft Computing, unpublished Ph.D. Dissertation, Royal Institute of Technology, Stockholm Sweden (2002)
[29] Klir, G., Folger, T.: Fuzzy Sets, Uncertainty, and Information. Prentice-Hall, NY (1988)
[30] Kulik, L., Eschenbach, C., Habel, C., Schmidtke, H.: A graded approach to directions between extended objects. In: Egenhofer, M.J., Mark, D.M. (eds.) GIScience 2002. LNCS, vol. 2478, pp. 119–131. Springer, Heidelberg (2002)
[31] Ladner, R., Petry, F., Cobb, M.: Fuzzy Set Approaches to Spatial Data Mining of Association Rules. Transactions on GIS 7#1, 123–138 (2003)
[32] Li, Y., Li, S.: A fuzzy topology for computing the interior, boundary, and exterior of spatial objects quantitatively in GIS. In: Shi, W., Liu, K. (eds.) Computers and Geosciences, vol. 33(7), pp. 898–915 (2007)
[33] MacMillan, R., Pettapiece, W., Nolan, S., Goddard, T.: A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. Fuzzy Sets and Systems 113, 81–109 (2000)
[34] Matsakis, P.: Understanding the Spatial Organization of Image Regions by Force Histograms. In: Matsakis, P., Sztandra, L. (eds.) Applying Soft Computing in Defining Spatial Relations, pp. 1–16. Physica Verlag, Heidelberg (2002)
[35] Morris, A.: A framework for modeling uncertainty in spatial databases. Transactions in GIS 7, 83–101 (2003)

[36] Papadias, D., Karacapilidis, N., Arkoumnais, D.: Processing Fuzzy Spatial Queries: A Configuration Similarity Approach. Int. J. Geographical Information Science 13, 93–128 (1999)
[37] Pawlak, Z.: Rough Sets. Int. Journal of Man-Machine Studies 21, 127–134 (1984)
[38] Pedrycz, W., Gomide, F.: An Introduction to Fuzzy Sets –Analysis and Design. MIT Press, Cambridge (1998)
[39] Robinson, V., Frank, A.: About Different Kinds of Uncertainty in Geographic Information Systems. In: Proceedings AUTOCARTO 7 Conference, pp. 440–449. American Society for Photogrammetry and Remote Sensing, Falls Church (1985)
[40] Roy, A., Stell, J.: Spatial Relations Between Indeterminate Regions. International Journal of Approximate Reasoning 27, 205–234 (2001)
[41] Schneider, M.: Spatial Data Types for Database Systems Doctoral Thesis, Fern Universität Hagen, GR (1995)
[42] Schockaert, S., De Cock, M., Cornelis, C., Kerre, E.: Fuzzy region connec-tion calculus: Representing vague topological information. International Journal of Approximate Reasoning 48, 314–331 (2008)
[43] Schockaert, S., De Cock, M., Cornelis, C., Kerre, E.: Fuzzy region connec-tion calculus: An interpretation based on closeness. International Journal of Approximate Reasoning 48, 332–347 (2008)
[44] Schockaert, S., De Cock, M., Kerre, E.: Spatial reasoning in a fuzzy region connection calculus. Artificial Intelligence 173(2), 258–298 (2009)
[45] Shi, X., Xing, A., Zhu, A., Wang, R.: Fuzzy Representation of Special Terrain Features Using a Similarity-based Approach. In: Petry, F., Robinson, V., Cobb, M. (eds.) Fuzzy Modeling with Spatial Information for Geographic Problems, pp. 233–252. Springer, Heidelberg (2005)
[46] Shi, W., Wang, S., Li, D., Wang, X.: Uncertainty-based Spatial Data Mining. In: Proceedings of Asia GIS Association, Wuhan, China, pp. 124–135 (2003)
[47] Skidmore, A.: Terrain position as mapped from a gridded digital elevation model. Int. J. Geographical Information Systems 4, 33–49 (1990)
[48] Somodevilla, M.: Fuzzy MBRs Modeling for Reasoning about Vague Regions. PhD thesis, Department of Electrical Engineering and Computer Science, Tulane University, New Orleans, LA (2003)
[49] Somodevilla, M., Petry, F.: Approximation of Topological Relations on Fuzzy Regions: An Approaah using Minimal Bounding Rectangles. In: Proceedings of the NAFIPS 2003 Conference, Chicago, vol. Il, pp. 1–23 (2003)
[50] Somodevilla, M., Petry, F.: Fuzzy Minimum Bounding Rectangles. In: DeCaluwe, R., DeTre, G., Bordogna, G. (eds.) Flexible Querying and Reasoning in Spatio-temporal Databases. Theory and Applications. Studies in Fuzziness and Soft Computing, pp. 237–264. Springer, Heidelberg (2004)
[51] Stoms, D.: Reasoning with Uncertainty in Intelligent Geographic Information Systems. In: Proceedings GIS 1987 2nd Annual International Conference on Geographic Information Systems, pp. 693–699. American Society for Photogrammetry and Remote Sensing, Falls Church (1987)
[52] Tang, X., Fang, Y., Kainz, W.: Fuzzy Topological Relations Between Fuzzy Spatial Objects. In: Wang, L., et al. (eds.) FSKD 2006. LNCS (LNAI), vol. 4223, pp. 324–333. Springer, Heidelberg (2006)
[53] Vazirgiannis, M.: Uncertainty Handling in Spatial Relationships. In: ACM-SAC 2000 Proceedings, Como Italy, pp. 215–221 (2000)

[54] Verstraete, J., De Tré, G., De Caluwe, R., Hallez, A.: Field Based Methods for the Modeling of Fuzzy Spatial Data. In: Petry, F., Robinson, V., Cobb, M. (eds.) Fuzzy Modeling with Spatial Information for Geographic Problems, pp. 41–70. Springer, Heidelberg (2005)

[55] Wang, F.: A Fuzzy Grammar and Possibility Theory-based Natural Language User Interface for Spatial Queries. International Journal of Fuzzy Sets and Systems 113, 147–159 (2000)

[56] Wang, S., Li, D., Shi, W., Wang, X.: Rough Spatial Description. In: International Archives of Photogrammetry and Remote Sensing, vol. XXXII, Commission II, pp. 503–510 (2002)

[57] Wang, S., Yuan, H., Chen, G., Li, D., Shi, W.: Rough Spatial Interpretation. In: Tsumoto, S., Słowiński, R., Komorowski, J., Grzymała-Busse, J.W. (eds.) RSCTC 2004. LNCS (LNAI), vol. 3066, pp. 435–444. Springer, Heidelberg (2004)

[58] Worboys, M.: Imprecision in Finite Resolution Spatial Data. Geoinformatica 2(3), 257–280 (1998)

[59] Worboys, M.: Computation with imprecise geospatial data. Computers, Environment and Urban Systems 22(2), 85–106 (1998)

[60] Yager, R.: On a General Class of Fuzzy Connectives. Fuzzy Sets and Systems 3, 235–242 (1980)

[61] Zadeh, L.: Fuzzy Sets. Information and Control 8, 338–353 (1965)

[62] Zhan, F.B.: Approximate analysis of binary topological relations between geographic regions with indeterminate boundaries. Soft Computing 2(2), 28–34 (1998)

[63] Zhang, J., Goodchild, M.: Uncertainty in Geographical Information. Taylor & Francis, London (2002)